

## Equations for the Transfer of Neutral Molecules and Ionic Species from Water to Organic phases

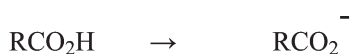
Michael H Abraham<sup>\*,†</sup> and William E Acree, Jr.<sup>‡</sup>

<sup>†</sup>*Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, U.K., and*

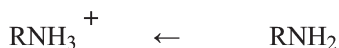
<sup>‡</sup>*Department of Chemistry, 1155 Union Circle Drive No. 305070, University of North Texas, Denton, Texas 76203-5017*

*m.h.abraham@ucl.ac.uk*

*Received November 17, 2009*



Hydrogen bond acid    Hydrogen bond base



Equations that can be used for partition coefficients of both neutral molecules and ions have been revised, and the term that is specific to ions has been re-evaluated. A new method has been devised for the determination of partition coefficients from water to organic solvents for carboxylate anions that is based on the variation of  $\text{p}K_a$  for carboxylic acids with solvent. Using these partition coefficients, descriptors for carboxylate anions, on the same scale as descriptors for neutral molecules, have been obtained for 71 such anions. For 13 carboxylate anions in a test set, descriptors could be predicted that in turn led to predictions of 78  $\log P$  values over six water–solvent systems with an absolute error of 0.13 and a standard deviation of only 0.55 log units. Descriptors have been obtained for 26 protonated amines as a training set and descriptors predicted for 17 protonated amines as a test set. The predicted descriptors in turn led to the prediction of 18  $\log P$  values for protonated amines with an absolute error of 0.09 and a standard deviation of 0.39 log units. The carboxylate anions are the strongest monofunctional hydrogen bond bases, and the protonated amines are the strongest monofunctional hydrogen bond acids that we have studied.

### Introduction

The partition of ionic species from water to organic phases is of ongoing interest. There are numerous systems that are models for biological processes, as well as biological processes themselves, in which ionic species and neutral molecules are transferred from water to organic phases such as membranes. Mälkiä et al.<sup>1</sup> have reviewed the permeation of drugs through biomembranes, including the relationship of partition coefficients for the neutral species and the ionized species in the case of ionizable solutes. However, the Gibbs energy of transfer of the ionic species was discussed<sup>1</sup> in terms

of variants of the Born equation<sup>2–5</sup> or of a purely empirical function of the surface electric field strength,<sup>6,7</sup> neither of which can easily be related to properties of the neutral species. Kedem and Katchalsky<sup>8</sup> deal also with permeation of ions and neutral molecules through membranes, this time in terms of transference numbers and frictional coefficients. How these are to be evaluated for drug molecules is not clear. In a very detailed analysis of partitioning across lipid bilayers, Mitragotri et al.<sup>9</sup> evaluated the work required to create a cavity in the lipid bilayer and the work required to create a diffusion path for the solute, using scaled particle theory, but restricted their analysis to nonelectrolytes.

(1) Mälkiä, A.; Murtomäki, L.; Urtti, A.; Kontturi, K. *Eur. J. Pharm. Sci.* **2004**, *23*, 13–47.

(2) Abraham, M. H.; Liszi, J. *J. Chem. Soc., Faraday Trans. 1* **1978**, *74*, 1604–1614.

(3) Abraham, M. H.; Liszi, J.; Meszaros, L. *J. Chem. Phys.* **1979**, *70*, 2491–2496.

(4) Abraham, M. H.; Liszi, J. *J. Chem. Soc., Faraday Trans. 1* **1980**, *76*, 1219–1231.

(5) Parsegian, A. *Nature* **1969**, *221*, 844–846.

(6) Osakai, T.; Ogata, A.; Ebina, K. *J. Phys. Chem. B* **1997**, *101*, 8341–8348.

(7) Osakai, T.; Ebina, K. *J. Phys. Chem. B* **1998**, *102*, 5691–5698.

(8) Kedem, O.; Katchalsky, A. *J. Gen. Physiol.* **1961**, *45*, 143–179.

(9) Mitragotri, S.; Johnson, M. E.; Blankschtein, D.; Langer, R. *Biophys. J.* **1999**, *77*, 1268–1283.

More recently, permeation through artificial membranes has been put forward for the prediction of biological processes, such as human intestinal absorption. Wohnsland and Faller,<sup>10</sup> in a novel study, showed that permeation from water through a hexadecane layer could be used to mimic intestinal absorption. They observed a peculiar dependence of permeability on the pH of the aqueous phase for ionizable drugs that they attributed to the presence of an unstirred water layer at the water/hexadecane interface.

Avdeef et al.<sup>11</sup> used a more complicated experimental method that involved experiments not only at different pH but also at different stirring rates. They were then able to dissect the overall observed permeability into contributions from the neutral (un-ionized) molecule and the ionized form. Thus, for the typical strong base atenolol,  $pK_a = 9.54$ , they found permeability coefficients in units of  $10^{-6} \text{ cm s}^{-1}$  of 32 for the neutral form and 1.1 for the ionized form for permeation through Caco cells and 191 for permeation of the ionized form through the unstirred water layer.

It would be of considerable interest if properties of neutral and ionized forms of acids and bases could be compared in some way. The variants of the Born equation are of little value, as they apply only to the ionized form, and solute descriptors calculated by the popular software methods CODESSA<sup>12</sup> and DRAGON<sup>13</sup> apply only to the neutral form. However, the COSMO-RS method has been used to calculate standard Gibbs energies of transfer from water to nitrobenzene and 4-(3-phenylpropyl)pyridine.<sup>14</sup> It is the aim of the present work to set out equations for transfer from water to organic phases that include both neutral species and ionic species. Of course, this requires that neutral species and ionic species are defined by the same types of descriptors. We have shown that neutral solutes can be described by five descriptors<sup>15,16</sup> and that equations containing only these five descriptors can be used to fit and to predict solute properties in a wide range of systems. These include partitions from water to some 40 solvents,<sup>17</sup> to ionic liquids,<sup>18</sup> and to aqueous biphasic systems,<sup>19</sup> as well as biological processes.<sup>20</sup> The general equation that we used is eq 1, where SP is a solute property in a given system. In the present work, SP will be a water to solvent partition coefficient, as log  $P$ .

$$SP = c + eE + sS + aA + bB + vV \quad (1)$$

The solute descriptors are as follows.<sup>15,16</sup>  $E$  is an excess molar refraction in  $\text{cm}^3 \text{ mol}^{-1}/10$ .  $S$  is a combined dipolarity/polarizability descriptor.  $A$  is the overall solute hydrogen bond acidity, and  $B$  is the overall solute hydrogen bond basicity.  $V$  is McGowan's characteristic molecular volume in  $\text{cm}^3 \text{ mol}^{-1}/100$ . The set of coefficients  $c$ ,  $e$ ,  $s$ ,  $a$ ,  $b$ , and  $v$  characterize the system and are determined by multiple linear regression analysis.

We used partition coefficients for ions from water to various solvents to assign the above descriptors to a number of ions<sup>21</sup> and showed that eq 1 could incorporate ions as well as neutral molecules if cations were assigned an additional descriptor  $J^+$  and anions an additional descriptor  $J^-$  to yield eq 2.

$$SP = c + eE + sS + aA + bB + vV + j^+J^+ + j^-J^- \quad (2)$$

Subsequently, we showed that eq 2 could be applied to partition of ions and neutral molecules from water to *o*-nitrophenyl octyl ether (NPOE).<sup>22</sup> Partition coefficients,  $P_{\text{npoe}}$ , for a wide range of nonelectrolytes had been determined by Carrupt et al.,<sup>23</sup> and a regression of log  $P_{\text{npoe}}$  against the solute descriptors yielded eq 3 for the neutral solutes.<sup>22</sup>

$$\log P_{\text{npoe}} = 0.121 + 0.600E - 0.459S - 2.246A - 3.879B + 3.574V \quad (3)$$

$$N = 55, R^2 = 0.935, SD = 0.308, F = 142$$

In eq 3 and elsewhere,  $N$  is the number of data points or compounds,  $R$  is the correlation coefficient,  $SD$  is the regression standard deviation, and  $F$  is the F-statistic. When values for 15 ions were included, eq 4 was obtained, with an  $SD$  value for the added 15 ions of 0.72 log units. This was the first time that a partition equation had been constructed that included data for both neutral species and ions.

$$\log P_{\text{npoe}} = 0.121 + 0.600E - 0.459S - 2.246A - 3.879B + 3.574V - 2.271J^+ + 0.432J^- \quad (4)$$

Since then, we have obtained equations for the partition of neutral species from water to ketones<sup>24</sup> and amide,<sup>25</sup> and have updated equations for partition of neutral species to chloroalkanes<sup>26</sup> and to alcohols,<sup>27</sup> and we aim to use these new equations to develop equations of the form of eq 2 that include ionic species as well as neutral species.

## Results and Discussion

The coefficients in the equations we use are collected in Table 1,<sup>17,24–27</sup> all obtained from data on neutral solutes.

As described previously, we use partition coefficients for ions based on the assumption that log  $P$  values for  $\text{Ph}_4\text{As}^+$ ,  $\text{Ph}_4\text{P}^+ = \log P$  values for  $\text{Ph}_4\text{B}^-$ . Log  $P$  values for single ions on the molar scale at 25 °C are in Table S1 (Supporting

- (10) Wohnsland, F.; Faller, B. *J. Med. Chem.* **2001**, *44*, 923–930.  
 (11) Avdeef, A.; Artursson, P.; Neuhoff, S.; Lazorova, L.; Gräsjo, J.; Tavelin, S. *Eur. J. Pharm. Sci.* **2005**, *24*, 333–349.  
 (12) Katritzky, A. R.; Lovanov, V. S.; Karelson, M. *CODESSA, Reference Manual V 2.13*; Semichem and the University of Florida: Gainesville, 1997.  
 (13) Tetko, I. V.; Gasteiger, J.; Todeschini, R.; Mauri, A.; Livingstone, D.; Ertl, P.; Palyulin, V. A.; Radchenko, E. V.; Zefirov, N. S.; Makarenko, A. S.; Tanchuk, V. Y.; Prokopenko, V. V. *J. Comput.-Aided Mol. Des.* **2005**, *19*, 453–63.  
 (14) MacDonald, S. M.; Opallo, M.; Klamt, A.; Eckert, F.; Marken, F. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3925–3933.  
 (15) Abraham, M. H. *Chem. Soc. Rev.* **1993**, *22*, 73–83.  
 (16) Abraham, M. H.; Ibrahim, A.; Zissimos, A. M. *J. Chromatogr. A* **2004**, *1037*, 29–47.  
 (17) Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R.; Acree, W. E., Jr. *J. Pharm. Sci.* Early view. DOI: 10.1002/jps.21922.  
 (18) Sprunger, L. M.; Gibbs, J.; Proctor, A.; Acree, W. E., Jr.; Abraham, M. H.; Meng, Y.; Yao, C.; Anderson, J. L. *Ind. Eng. Chem.* **2009**, *48*, 4145–4154.  
 (19) Abraham, M. H.; Zissimos, A. M.; Huddleston, J. G.; Willauer, H. D.; Rodgers, R. D.; Acree, W. E., Jr. *Ind. Eng. Chem. Res.* **2003**, *42*, 413–418.  
 (20) Zhao, Y. H.; Le, J.; Abraham, M. H.; Hersey, A.; Eddershaw, P. E.; Luscombe, C. N.; Butina, D.; Beck, G.; Sherborne, B.; Cooper, I.; Platts, J. A. *J. Pharm. Sci.* **2001**, *90*, 749–784.

- (21) Abraham, M. H.; Zhao, Y. H. *J. Org. Chem.* **2004**, *69*, 4677–4685.  
 (22) Abraham, M. H.; Zhao, Y. H. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2418–2422.  
 (23) Liu, X.; Bouchard, G.; Muller, N.; Galland, A.; Girault, H.; Testa, B.; Carrupt, P.-A. *Helv. Chim. Acta* **2003**, *86*, 3533–3547.  
 (24) Abraham, M. H.; Acree, W. E., Jr.; Leo, A. J.; Hoekman, D. *New J. Chem.* **2009**, *33*, 568–573.  
 (25) Abraham, M. H.; Acree, W. E., Jr.; Cometto-Muñiz, J. E. *New J. Chem.* **2009**, *33*, 2034–2043.  
 (26) Sprunger, L. M.; Achi, S. S.; Acree, W. E., Jr.; Abraham, M. H.; Leo, A. J.; Hoekman, D. *Fluid Phase Eq.* **2009**, *281*, 144–162.  
 (27) Sprunger, L. M.; Achi, S. S.; Pointer, R.; Blake-Taylor, B. H.; Acree, W. E., Jr.; Abraham, M. H. *Fluid Phase Eq.* **2009**, *286*, 170–174.

TABLE 1. Coefficients in eq 1 for the Partition of Neutral Solutes from Water to Organic Phases (SP = log *P* at 25 °C)

solvent	<i>c</i>	<i>e</i>	<i>s</i>	<i>a</i>	<i>b</i>	<i>v</i>
1,2-dichloroethane, 12DCE	0.183	0.294	-0.134	-2.801	-4.291	4.180
dichloromethane, DCM	0.319	0.102	-0.187	-3.058	-4.090	4.324
NPOE <sup>a</sup>	0.121	0.600	-0.459	-2.246	-3.879	3.574
propanone	0.313	0.312	-0.121	-0.608	-4.753	3.942
acetonitrile, MeCN	0.413	0.077	0.326	-1.566	-4.391	3.364
nitromethane	0.023	-0.091	0.793	-1.463	-4.364	3.460
nitrobenzene	-0.152	0.525	0.081	-2.332	-4.494	4.187
<i>N</i> -methylpyrrolidinone, NMP	0.147	0.532	0.275	0.840	-4.794	3.674
dimethylformamide, DMF	-0.305	-0.058	0.343	0.357	-4.865	4.486
dimethylacetamide, DMA	-0.271	0.084	0.209	0.915	-5.003	4.557
dimethylsulfoxide, DMSO	-0.194	0.327	0.791	1.260	-4.540	3.361
methanol	0.276	0.334	-0.714	0.243	-3.320	3.549
ethanol	0.222	0.471	-1.035	0.326	-3.596	3.857
propan-1-ol	0.139	0.405	-1.029	0.247	-3.767	3.986
butan-1-ol	0.165	0.401	-1.011	0.056	-3.958	4.044
hexan-1-ol	0.115	0.492	-1.164	0.054	-3.971	4.131
propan-2-ol	0.099	0.344	-1.049	0.406	-3.827	4.033
<i>tert</i> -butyl alcohol	0.211	0.171	-0.947	0.331	-4.085	4.109
wet octan-1-ol	0.088	0.562	-1.054	0.034	-3.460	3.814
formamide	-0.171	0.070	0.308	0.589	-3.152	2.432

<sup>a</sup>*o*-Nitrophenyl octyl ether.

Information). Most of them are those listed previously,<sup>21,22</sup> with some alterations or additions from the valuable compilations of Marcus et al.<sup>28–30</sup> Although we listed<sup>21,22</sup> descriptors for many of these ions, we have recalculated descriptors using the most recent equations. Previously, we took<sup>21,22</sup>  $j^+$  as  $-3$  for all solvents and  $j^-$  as  $0$  for aprotic solvents and as  $3$  for the alcohol solvents. Now that we have more data, we have no need to fix  $j^+$  and  $j^-$ , and so we allowed these coefficients to float. Since we fix the coefficients  $e$ ,  $s$ ,  $a$ ,  $b$ , and  $v$  to those for the solvent equations for the neutral solutes, the coefficients  $j^+$  and  $j^-$ , are the only extra coefficients in the equations that are used to deal with the various cations and anions. Details of the new descriptors for the ions in Table S1 (Supporting Information) are given in Table 2, together with the number of solvent systems used and the standard deviation in the experimental and fitted log *P* values for the ions. The two extra solvent coefficients are given in Table 3, together with the number of ions used to obtain the coefficients and the standard deviation in the experimental and fitted log *P* values for the ions. In order to make calculations for organic ions more simple, we took  $E_i$  for the tetraalkylammonium ions as  $-0.10$ , quite close to the corresponding amines, and calculated  $V_i$  by the usual McGowan procedure, using the molecular formula for the ions. We denote descriptors for ions by the subscript “*i*” but note that descriptors for ions and neutral molecules are on the same scale.

The two ionic descriptors in Table 3 are quite close to those we obtained before, but now we have an estimate of how good the fit is between experimental and calculated log *P* values. With just one additional descriptor for anions and one for cations, the log *P* values can be fitted to about 0.25 log unit. It is not easy to assess the experimental error in the log *P* values, but for transfer to methanol, five determinations of  $K^+$  differ by 0.80 log units<sup>29</sup> and three determinations of  $F^-$  differ by 0.70 log units.<sup>30</sup> Occasionally, there are much larger discrepancies; log *P* for transfer of  $F^-$  to acetonitrile is given as 12.44<sup>28</sup> and 7.88<sup>30</sup> log units. Of course, we have had to make some selection as regards

TABLE 2. Descriptors for Ions, the Number of Solvents Used, and the Standard Deviation in the Log *P* Values

ion	$E_i$	$S_i$	$A_i$	$B_i$	$V_i$	$J^+$	$J^-$	<i>N</i>	SD
$Li^+$	-0.02	2.11	1.30	0.00	0.014	0.154	0	7	0.284
$Na^+$	-0.02	2.31	1.22	0.00	0.033	0.316	0	8	0.258
$K^+$	0.00	2.57	1.21	0.00	0.092	0.357	0	11	0.249
$Rb^+$	0.02	2.55	1.05	0.00	0.130	0.477	0	10	0.315
$Cs^+$	0.10	2.60	1.17	0.00	0.177	0.438	0	11	0.203
$NMe_4^+$	-0.10	1.31	0.68	0.00	0.7635	1.235	0	10	0.120
$NEt_4^+$	-0.10	1.85	0.51	0.00	1.3571	1.475	0	11	0.153
$NPr_4^+$	-0.10	2.02	0.42	0.00	1.9207	1.552	0	8	0.106
$NBu_4^+$	-0.10	2.82	0.61	0.00	2.4843	1.418	0	6	0.123
$Ph_4P^+$	2.22	3.11	0.04	0.92	2.766	0.480	0	10	0.163
$Ph_4As^+$	2.22	3.20	0.07	0.91	2.811	0.581	0	13	0.172
$F^-$	-0.05	3.76	0.00	2.42	0.105	0	2.385	5	0.285
$Cl^-$	0.10	3.52	0.00	2.32	0.228	0	2.363	15	0.239
$Br^-$	0.17	2.74	0.00	1.82	0.307	0	1.567	13	0.145
$I^-$	0.38	3.55	0.00	1.34	0.408	0	1.251	12	0.194
$ClO_4^-$	-0.16	5.14	0.00	0.99	0.493	0	1.290	7	0.172
$CN^-$	0.07	1.99	0.00	1.86	0.231	0	1.630	4	0.083
$NO_3^-$	0.17	1.98	0.00	1.97	0.320	0	1.703	6	0.212
$SCN^-$	0.40	3.38	0.00	1.24	0.365	0	1.242	8	0.110
$N_3^-$	0.16	3.04	0.00	1.82	0.282	0	1.718	8	0.158
$Ph_4B^-$	1.95	2.72	0.18	1.15	2.700	0	-0.188	15	0.417

the experimental values that we used, but an average SD value of about 0.25 log units seems very reasonable.

About the same standard deviation is observed in fits with respect to the equations, Table 3. However, we encountered difficulties with a number of aprotic solvents and were unable to fit the log *P* values for cations satisfactorily. In the case of nitrobenzene,<sup>28–30</sup> we were unable to fit the cations, and the fit for the anions (see Table 3) is so poor as to be of little value. This is very unfortunate because nitrobenzene has been one of the most studied solvents as regards partition of ions. Similarly, we were unable to obtain any fits to data in solvent dichloromethane.<sup>31</sup> We have no explanation of why transfers of cations to 1,2-dichloroethane, NPOE, propanone, acetonitrile, *N*-methylpyrrolidinone, and DMSO can be fitted satisfactorily but transfers to nitromethane, nitrobenzene, dimethylformamide, and dimethylacetamide cannot.

(28) Marcus, Y. *Pure Appl. Chem.* **1983**, *55*, 977–1021.(29) Kalidas, C.; Heffer, G.; Marcus, Y. *Chem. Rev.* **2000**, *100*, 819–852.(30) Marcus, Y. *Chem. Rev.* **2007**, *107*, 3880–3897.(31) de Namor, A. F. D.; Traboulssi, R.; Salazar, F. F.; de Acosta, V. D.; de Vizcardo, Y. F.; Portugal, J. M. *J. Chem. Soc., Faraday Trans. 1* **1989**, *85*, 2705–2712.

**TABLE 3. Coefficients for Solvents, the Number of Ions Used, and the Standard Deviation in the Log  $P$  Values**

solvent	$j^+$	$j^-$	$N$	SD
1,2-dichloroethane	-3.429	-0.025	14	0.193
NPOE	-2.314	0.350	10	0.274
propanone	-2.288	0.078	14	0.288
acetonitrile	-2.243	0.101	17	0.202
nitromethane <sup>a</sup>		-0.149	5	0.286
nitrobenzene <sup>a</sup>		1.134	6	0.408
<i>N</i> -methylpyrrolidinone	-1.797	0.105	9	0.251
dimethylformamide <sup>a</sup>		0.415	9	0.191
dimethylacetamide <sup>a</sup>		0.286	5	0.248
DMSO	-3.387	0.132	22	0.156
methanol	-2.609	3.027	22	0.198
ethanol	-3.170	3.085	19	0.156
propan-1-ol	-3.077	2.834	17	0.197
butan-1-ol	-3.605	2.685	15	0.287
hexan-1-ol	-3.100	2.940	6	0.331
propan-2-ol	-3.896	2.889	12	0.228
<i>tert</i> -butyl alcohol	-4.455	2.953	8	0.233
formamide	-3.152	2.432	12	0.238

<sup>a</sup>Anions only.

**Monocarboxylate anions.** There have been a few studies in which Gibbs energies of transfer of carboxylate anions from water to an organic phase, equivalent to log  $P$  values, have been determined through the use of various redox systems. These include the work of MacDonald et al.<sup>14</sup> on the 4-(3-phenylpropyl)pyridine solvent phase, of Komorsky-Lovrić et al.<sup>32</sup> on solvent nitrobenzene, and of Bouchard et al. on solvents 1,2-dichloroethane<sup>33</sup> and octan-1-ol.<sup>34</sup> As well as the electrochemical redox method, the most usual method of obtaining log  $P$  values is through measurements of solubility, corrected for ion-pair association and activity coefficients. They can also be obtained through a thermodynamic cycle, although to our knowledge this has never previously been used. Consider eqs 5 and 6 for ionization in water and in some other phase.



Then for transfer from water to the other phase, log  $P(\text{A}^-)$  is given by eq 7

$$\log P(\text{A}^-) = \log P(\text{HA}) - \log P(\text{H}^+) + \text{p}K_{\text{a}}(\text{aq}) - \text{p}K_{\text{a}}(\text{s}) \quad (7)$$

where  $P(\text{HA})$  is the corresponding partition coefficient for the neutral carboxylic acid. This can, in principle, be obtained through solubility measurements of the neutral carboxylic acid in water and the other phase, but for phases that are organic solvents it can easily be estimated through the equations given in Table 1 and the descriptors for the neutral compound. Then  $\text{p}K_{\text{a}}(\text{aq})$ ,  $\text{p}K_{\text{a}}(\text{s})$ , and crucially, log  $P(\text{H}^+)$  are required in order to calculate log  $P(\text{A}^-)$ . Collections of  $\text{p}K_{\text{a}}(\text{aq})$  and  $\text{p}K_{\text{a}}(\text{s})$  for benzoic acids<sup>35</sup> and aliphatic car-

(32) Komorsky-Lovrić, S.; Riedl, K.; Gulaboski, R.; Mirčeski, V.; Scholz, F. *Langmuir* **2002**, *18*, 8000–8005.

(33) Bouchard, G.; Carrupt, P. A.; Testa, B.; Gobry, V.; Girault, H. H. *Chem.—Eur. J.* **2002**, *8*, 3478–3484.

(34) Bouchard, G.; Galland, A.; Carrupt, P. A.; Gulaboski, R.; Girault, H. H. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3748–3751.

(35) Jover, J.; Bosque, R.; Sales, J. *QSAR Comb. Sci.* **2008**, *27*, 563–581.

(36) Jover, J.; Bosque, R.; Sales, J. *QSAR Comb. Sci.* **2008**, *27*, 1204–1215.

**TABLE 4. Values of log  $P(\text{H}^+)$  for Transfer from Water to Organic Phases, Molar Scale at 25 °C**

solvent	log $P(\text{H}^+)$				
	ref 28	ref 29	ref 40 <sup>a</sup>	this work	taken
propanone				-2.56	-2.56
acetonitrile	-8.12	-7.85	-8.02		-7.85
nitromethane				-0.99	-0.99
<i>N</i> -methylpyrrolidinone	4.38			3.33	3.33
dimethylformamide	3.15	2.52			2.52
dimethylacetamide				4.90	4.90
DMSO	3.40	3.40			3.40
methanol	-1.82	-1.52			-1.82
ethanol	-1.94	-1.94	-1.72	-1.60	-1.60
propan-1-ol	-1.58		-1.52	1.15	1.15
propan-2-ol			-0.42	-1.93	-1.93
<i>tert</i> -butyl alcohol			-1.02	-3.30	-3.30
hexan-1-ol			-2.22		-2.22

<sup>a</sup>Relative to log  $P(\text{H}^+) = -1.82$  in methanol.

boxylic acids<sup>36</sup> have been published by Jover et al., and Roses et al. have compiled a comprehensive list of  $\text{p}K_{\text{a}}(\text{s})$  in methanol<sup>37</sup> and a few in *tert*-butyl alcohol.<sup>38</sup> Chantooni and Kolthoff<sup>39–42</sup> have determined a number of other values, especially for the alcohols propan-2-ol, *tert*-butyl alcohol, and hexan-1-ol, and Ding et al.<sup>43</sup> have listed values in DMSO and acetonitrile. Pytela et al.<sup>44–48</sup> have determined  $\text{p}K_{\text{a}}(\text{s})$  for a very large number of benzoic acids in various solvents, and we have also taken a number of values from the compilation of Izutsu.<sup>49</sup> We used also values for a few 3,4-disubstituted benzoic acids from Pytela et al.<sup>50</sup> and saved the rest to use as a test set. A number of log  $P(\text{H}^+)$  values are listed by Marcus,<sup>28</sup> by Kalidas et al.,<sup>29</sup> and by Chantooni and Kolthoff<sup>40</sup> as shown in Table 4. Of course, in order to be compatible with log  $P$  values for the ions in Table S1 (Supporting Information), log  $P(\text{H}^+)$  must also be on the  $\text{Ph}_4\text{As}^+$ ,  $\text{Ph}_4\text{P}^+ = \text{Ph}_4\text{B}^-$  scale.

As an example, we show a calculation for the acetate ion in Table 5, using the  $\text{p}K_{\text{a}}$  data given by Jover et al.<sup>35</sup> We can fit the obtained log  $P$  values for the acetate ion quite well, as shown in Table 5, where the standard deviation between fitted and observed values, SD = 0.40 log units for 8 data points. However, the fit for log  $P$  values determined by the solubility method was so poor as to be of little value,

(37) Rived, F.; Roses, M.; Bosch, E. *Anal. Chim. Acta* **1998**, *374*, 309–324.

(38) Bosch, E.; Roses, M. *Talanta* **1989**, *36*, 627–632.

(39) (a) Chantooni, M. K., Jr.; Kolthoff, I. M. *J. Phys. Chem.* **1973**, *77*, 527–533. (b) Chantooni, M. K., Jr.; Kolthoff, I. M. *J. Phys. Chem.* **1976**, *79*, 1176–1182.

(40) Chantooni, M. K., Jr.; Kolthoff, I. M. *J. Phys. Chem.* **1978**, *82*, 994–1000.

(41) (a) Chantooni, M. K., Jr.; Kolthoff, I. M. *Anal. Chem.* **1978**, *50*, 1440–1446. (b) Chantooni, M. K., Jr.; Kolthoff, I. M. *Anal. Chem.* **1979**, *51*, 133–140.

(42) Kolthoff, I. M.; Chantooni, M. K., Jr. *J. Am. Chem. Soc.* **1976**, *24*, 7465–7470.

(43) Ding, F.; Smith, J. M.; Wang, H. *J. Org. Chem.* **2009**, *74*, 2679–2691.

(44) Ludwig, M.; Baron, V.; Kalfus, K.; Pytela, O.; Večeřa, M. *Collect. Czech. Chem. Commun.* **1986**, *51*, 2135–2142.

(45) Pytela, O.; Kulhánek, J.; Ludwig, M. *Collect. Czech. Chem. Commun.* **1994**, *59*, 1637–1644.

(46) Kulhánek, J.; Pytela, O. *Collect. Czech. Chem. Commun.* **1995**, *60*, 829–840.

(47) Kulhánek, J.; Pytela, O. *Collect. Czech. Chem. Commun.* **1997**, *62*, 913–924.

(48) Pytela, O.; Kulhánek, J. *Collect. Czech. Chem. Commun.* **2002**, *67*, 596–608.

(49) Izutsu, K. *Acid-base dissociation constants in dipolar aprotic solvents*; IUPAC Chemical Data Series No. 35; Blackwell Scientific Publications: Oxford, 1990.

(50) Pytela, O.; Kulhánek, J.; Ludwig, M.; Riha, V. *Collect. Czech. Chem. Commun.* **1994**, *59*, 627–638.

SD = 2.20 log units for 9 data points. We therefore use the  $pK_a$  method to obtain log  $P$  values for the ionized forms of a number of carboxylic acids.

For solvents where  $pK_a$  values for the carboxylic acids are available, but not log  $P(H^+)$ , it is possible to obtain the corresponding log  $P(RCO_2^-)$  values if log  $P(H^+)$  can be evaluated by a trial-and-error method. For the (dry) solvents propan-2-ol, *tert*-butyl alcohol, nitromethane, and *N*-methylpyrrolidinone (NMP) there was enough data to obtain values

**TABLE 5. Calculation of log  $P(A^-)$  for the Acetate Ion Using the Thermodynamic Cycle of eq 5 and eq 6 and Fitted Values Using Descriptors for the Acetate Ion<sup>a</sup>**

phase	$pK_a(aq)$	$pK_a(s)$	log $P(H^+)$	log $P(HA)$	log $P(A^-)$	log $P(A^-)$ , fit
methanol	4.75	9.70	-1.82	0.25	-2.88	-3.02
ethanol	4.75	10.30	-1.60	0.10	-3.85	-4.27
propan-2-ol	4.75	11.30	-1.93	-0.83	-5.45	-4.90
<i>tert</i> -butyl alcohol	4.75	14.20	-3.30	-0.68	-6.83	-6.90
acetonitrile	4.75	22.30	-7.85	-0.70	-10.40	-10.01
DMA	4.75	12.60	4.90	0.37	-12.38	-11.82
DMF	4.75	13.25	2.52	0.07	-10.95	-10.98
DMSO	4.75	12.30	3.40	1.05	-9.90	-9.88
NMP	4.75	13.30	3.33	0.55	-11.33	-11.23
nitromethane	4.75	14.41	-0.99	-0.71	-9.38	-9.84
propanone	4.75	18.33	-2.56	-0.32	-11.34	-11.84

<sup>a</sup>Using the descriptors for the acetate ion in Table 6.

of log  $P(H^+)$  as shown in Table 4. Our value for propan-1-ol, 1.15, seems quite out of line by comparison to values for methanol and ethanol and is far away from the literature value<sup>40</sup> of -1.52 log units. However, our value is the average of values obtained for 12 carboxylic acids, with a standard deviation of no more than 0.13 log units.

**Calculations for Carboxylate Monoanions.** We used the log  $P(H^+)$  values in the last column of Table 4 and reported values of  $pK_a(aq)$  and  $pK_a(s)$ .<sup>35-49</sup> Log  $P$  for the neutral carboxylic acid was calculated from the descriptors for the acid and the coefficients given in Table 1. The data used for each acid are given in full in the Supporting Information (Table S2), and in Table 6 are the obtained descriptors for 71 carboxylate anions, together with the number of solvents used, and SD for the log  $P(RCO_2^-)$  values.

In our previous work,<sup>21</sup> we gave rather complicated methods for estimating  $V_i$  and  $E_i$  for ionic species. We denote ionic descriptors by the subscript "i" but note that the descriptors are on the same scale as those for neutral compounds. Many drug molecules contain the ionizable  $-CO_2H$  group, and it would be convenient if there was an easy method to estimate  $V_i$  and  $E_i$ . We find that  $V_i$  can simply be calculated by McGowan's method from the molecular formula of the ionic species (that is  $C_2H_3O_2$  for the acetate ion) and that  $E_i$  for the carboxyl group anion can be taken as slightly more than  $E$  for the corresponding carboxylic acid,

**TABLE 6. Descriptors for Carboxylate Anions Obtained Using the  $pK_a$  Method**

anion	$E_i$	$S_i$	$A_i$	$B_i$	$V_i$	$\mathcal{F}$	N	SD	anion	$E_i$	$S_i$	$A_i$	$B_i$	$V_i$	$\mathcal{F}$	N	SD
acetate	0.415	2.19	0.00	2.93	0.4433	2.075	12	0.362	4-hydroxybenzoate	1.080	3.78	0.04	3.05	0.9689	2.382	7	0.284
propanoate	0.383	1.69	0.00	2.91	0.5842	1.822	9	0.261	2-cyanobenzoate	1.010	4.37	0.00	2.89	1.0649	2.315	4	0.036
butanoate	0.360	2.18	0.00	3.05	0.7251	2.020	4	0.222	3-cyanobenzoate	1.010	3.70	0.00	2.82	1.0649	2.192	6	0.167
isobutanoate	0.350	1.61	0.00	2.97	0.7251	1.992	6	0.199	4-cyanobenzoate	1.010	3.96	0.00	2.78	1.0649	2.248	6	0.183
pentanoic	0.355	2.15	0.00	3.04	0.8660	2.082	4	0.341	2-nitrobenzoate	1.140	3.92	0.00	2.96	1.0844	2.132	6	0.099
chloroacetate	0.577	2.44	0.00	2.56	0.5657	1.860	10	0.218	3-nitrobenzoate	1.140	3.60	0.00	2.79	1.0844	2.217	8	0.236
dichloroacetate	0.632	2.53	0.00	2.18	0.6881	1.426	9	0.171	4-nitrobenzoate	1.140	3.50	0.00	2.79	1.0844	2.212	9	0.186
cyanoacetate	0.530	3.32	0.00	2.78	0.5980	1.947	5	0.127	4-aminobenzoate	1.225	4.07	0.25	3.26	1.0100	2.357	8	0.156
2,3-dibromopropanoate	0.867	3.49	0.00	2.29	0.9342	1.582	5	0.154	3-NHCOMe-benzoate	1.270	3.53	0.31	3.34	1.3075	2.188	5	0.076
hydroxyacetate	0.484	2.45	0.01	2.67	0.5020	1.689	5	0.033	4-NHCOMe-benzoate	1.270	3.68	0.24	3.35	1.3075	2.244	5	0.145
monomethylsuccinate	0.430	2.26	0.00	3.20	0.9404	1.882	6	0.086	3-SH-benzoate	1.370	3.36	0.01	2.77	1.0737	2.227	5	0.178
monomethylmaleate	0.630	2.73	0.00	3.18	0.8974	1.897	5	0.277	4-SH-benzoate	1.370	3.44	0.00	2.85	1.0737	2.191	5	0.094
benzoate	0.880	3.64	0.00	2.88	0.9102	2.395	13	0.257	2-MeSO-benzoate	1.440	4.11	0.00	3.36	1.2733	2.085	4	0.054
3-methylbenzoate	0.880	3.25	0.00	2.85	1.0511	2.228	8	0.145	4-MeSO-benzoate	1.440	3.92	0.00	3.48	1.2733	2.072	4	0.107
4-methylbenzoate	0.880	3.10	0.00	2.88	1.0511	2.167	7	0.180	2-MeSO <sub>2</sub> -benzoate	1.210	4.59	0.00	3.40	1.3320	2.252	4	0.057
4- <i>tert</i> -butylbenzoate	0.880	3.16	0.00	2.88	1.4738	2.229	5	0.193	3-MeSO <sub>2</sub> -benzoate	1.210	4.00	0.00	3.23	1.3320	2.121	5	0.171
3-fluorobenzoate	0.952	2.96	0.00	2.73	0.9279	2.175	5	0.141	4-MeSO <sub>2</sub> -benzoate	1.210	4.15	0.00	3.26	1.3320	2.188	5	0.162
4-fluorobenzoate	0.952	2.82	0.00	2.82	0.9279	2.174	5	0.192	3-NH <sub>2</sub> SO <sub>2</sub> -benzoate	1.500	3.99	0.04	3.39	1.2909	2.205	5	0.148
3-chlorobenzoate	0.990	3.13	0.00	2.57	1.0326	2.034	9	0.102	4-NH <sub>2</sub> SO <sub>2</sub> -benzoate	1.500	3.65	0.15	3.38	1.2909	2.102	5	0.059
4-chlorobenzoate	0.990	3.37	0.00	2.60	1.0326	2.179	8	0.151	2,6-difluorobenzoate	0.839	3.87	0.00	2.92	0.9456	2.468	9	0.170
3-bromobenzoate	1.150	3.47	0.00	2.56	1.0852	2.197	8	0.206	2,6-dichlorobenzoate	1.100	4.77	0.00	2.95	1.1550	2.624	8	0.118
4-bromobenzoate	1.150	3.25	0.00	2.60	1.0852	2.148	8	0.163	2,6-dibromobenzoate	1.540	4.97	0.00	3.07	1.2602	2.712	8	0.117
3-iodobenzoate	1.460	3.31	0.00	2.57	1.1684	2.082	7	0.270	2,6-diiodobenzoate	2.160	5.28	0.00	3.22	1.4266	2.678	9	0.223
4-iodobenzoate	1.460	3.29	0.00	2.61	1.1684	2.059	5	0.207	2,6-dimethylbenzoate	0.880	3.08	0.00	2.97	1.1920	2.259	9	0.107
3-trifluoromethylbenzoate	0.450	2.99	0.00	2.52	1.1042	2.119	5	0.141	2,6-dimethoxybenzoate	1.030	4.27	0.00	3.51	1.3094	2.527	8	0.125
3-methoxybenzoate	0.980	3.32	0.00	3.00	1.1098	2.254	6	0.169	2,6-diethoxybenzoate	1.020	4.09	0.00	3.50	1.5912	2.377	9	0.202
4-methoxybenzoate	1.050	4.00	0.00	3.05	1.1098	2.400	8	0.189	2,6-dipropoxybenzoate	1.010	4.12	0.00	3.52	1.8730	2.349	9	0.186
3-phenoxybenzoate	1.540	3.54	0.00	2.78	1.5767	2.088	5	0.147	2,6-diisopropoxybenzoate	0.990	4.30	0.00	3.65	1.8730	2.418	9	0.178
2-CHO-benzoate	1.160	3.96	0.00	2.75	1.0668	2.180	4	0.136	2,6-dibutoxybenzoate	1.000	4.31	0.00	3.51	2.1548	2.411	9	0.171
4-CHO-benzoate	1.160	3.64	0.00	2.83	1.0668	2.050	4	0.071	2,6-dinitrobenzoate	1.400	5.80	0.00	3.04	1.2586	2.566	9	0.327
2-MeCO-benzoate	1.090	3.92	0.00	3.16	1.2077	2.082	4	0.086	3,4-dimethylbenzoate	0.880	3.40	0.00	2.92	1.1920	2.251	7	0.142
3-MeCO-benzoate	1.090	3.66	0.00	3.04	1.2077	2.2010	6	0.172	3,4-dichlorobenzoate	1.100	3.04	0.00	2.49	1.1550	2.110	7	0.219
4-MeCO-benzoate	1.090	3.99	0.00	3.02	1.2077	2.280	6	0.203	3,5-dimethylbenzoate	0.880	3.44	0.00	2.94	1.1920	2.249	6	0.088
2-MeCO <sub>2</sub> -benzoate	1.000	4.15	0.00	3.28	1.2664	2.256	4	0.089	3,5-dinitrobenzoate	1.400	4.71	0.00	2.66	1.2586	2.005	5	0.048
4-MeCO <sub>2</sub> -benzoate	1.000	3.67	0.00	3.07	1.2664	2.164	4	0.089	3-nitro-4-chlorobenzoate	1.400	3.97	0.00	2.60	1.2068	2.119	6	0.261
3-hydroxybenzoate	1.060	3.65	0.13	3.07	0.9689	2.344	8	0.238									

**TABLE 7.** Comparison of Descriptors for the Carboxylate Anion, with those for Other Compounds

compd	<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>
acetate	0.415	2.19	0.00	2.93
benzoate	0.880	3.64	0.00	2.88
F <sup>-</sup>	-0.05	3.76	0.00	2.42
Cl <sup>-</sup>	0.10	3.52	0.00	2.32
triethylamine	1.01	0.15	0.00	0.79
<i>N,N</i> -dimethylacetamide <sup>a</sup>	0.363	1.38	0.00	0.80
succinamide	0.850	1.98	0.97	1.13
DMSO <sup>a</sup>	0.522	1.72	0.00	0.97
tributylphosphine oxide	0.00	1.15	0.00	1.64
diphenylphosphoric acid	1.50	1.55	0.81	1.45
benzamide	0.99	0.50	0.49	0.67
benzenesulfonamide	1.13	1.56	0.55	0.86
barbituric acid	1.09	1.19	0.49	1.16
atrazine	1.22	1.29	0.17	1.01

<sup>a</sup>Values for the compound as a solute.

by 0.15 unit. This reduces the number of descriptors that have to be determined to four, that is  $S_i$ ,  $A_i$ ,  $B_i$ , and  $J^-$ , or for the majority of anions with  $A_i = 0$ , to three. We used the trial-and-error procedure of ‘Solver’ in Microsoft Excel to obtain the combination of descriptors that gave the best fit to the observed  $\log P$  values for the anions. The best fit criterion was the standard deviation of the fitted and observed  $\log P$  values. In our experience, the method yields the same best fit descriptors, no matter what the starting values of the “unknown” descriptors are.

The main features of the descriptors for the ions are that  $S_i$  and  $B_i$  are very large, and, as expected, conversion of  $-\text{CO}_2\text{H}$  into  $-\text{CO}_2^-$  results in loss of all hydrogen bond acidity.

A comparison of values of  $S$ ,  $A$ , and  $B$  for some mono-functional compounds is shown in Table 7. It can be seen that the carboxylate ion is a very powerful hydrogen bond base, much more powerful than any neutral functional group and even more powerful than ions such as the fluoride or chloride ion. The carboxylate ion, especially if derived from benzoic acids, also has very large values of  $S$ , the polarizability/dipolarity descriptor.

There are enough data in Table 6 to attempt predictions of the ionic descriptors. Since we use the descriptors for the neutral carboxylic acids to obtain  $\log P(\text{HA})$  in eq 7, it was of interest to see if we could use the same descriptors to predict those for the corresponding ions. We find the following equations, where we include for completeness those for  $E_i$  and  $V_i$ .

$$E_i = 0.15 + 1.00E \quad (8)$$

$$S_i = 1.224 + 0.908E + 0.827S + 0.453V \quad (9)$$

$$A_i = -0.208 - 0.058S + 0.354A + 0.076A \quad (10)$$

$$B_i = 2.150 - 0.204S + 1.217B + 0.314V \quad (11)$$

$$V_i = -0.0215 + 1.00V \quad (12)$$

$$J^- = 1.793 + 0.267E - 0.195S + 0.350V \quad (13)$$

We can test the usefulness of these equations by predicting  $\log P$  for all 71 ions in Table 6 from water to a number of solvents and then comparing the obtained  $\log P$  values with the values obtained from the descriptors in Table 6. To do this, we use eqs 8–13 to predict descriptors for the 71 ions and then insert the predicted descriptors into eq 14–19 to

**TABLE 8.** Comparison of  $\log P$  Values for Carboxylate Anions Predicted through eqs 8–13 with Fitted Values Using Descriptors in Table 6

solvent	AE	AAE	RMSE	SD	<i>N</i>	avg $\log P^a$
methanol	0.00	0.21	0.26	0.26	71	-1.15
DMSO	-0.02	0.47	0.60	0.60	71	-6.34
DMF	-0.01	0.46	0.59	0.60	71	-7.57
propanone	-0.01	0.51	0.65	0.66	71	-9.27
acetonitrile	-0.02	0.44	0.57	0.58	71	-7.37
12DCE	-0.01	0.48	0.62	0.63	71	-8.12

<sup>a</sup>Average of the observed  $\log P$  value.

yield predicted  $\log P$  values for the carboxylate anions. The “experimental” descriptors in Table 6 yield  $\log P$  values that are “fitted” values.

$$\log P(\text{MeOH}) = 0.276 + 0.334E - 0.714S + 0.243A - 3.320B + 3.549V + 3.027J^- \quad (14)$$

$$\log P(\text{DMSO}) = -0.194 + 0.327E + 0.791S + 1.260A - 4.540B + 3.361V + 0.132J^- \quad (15)$$

$$\log P(\text{DMF}) = -0.305 - 0.058E + 0.343S + 0.357A - 4.865B + 4.486V + 0.415J^- \quad (16)$$

$$\log P(\text{propanone}) = 0.313 + 0.312E - 0.121S - 0.608A - 4.753B + 3.942V + 0.078J^- \quad (17)$$

$$\log P(\text{MeCN}) = 0.413 + 0.077E + 0.326S - 1.556A - 4.391B + 3.364V + 0.101J^- \quad (18)$$

$$\log P(\text{12DCE}) = 0.183 + 0.294E - 0.134S - 2.801A - 4.291B + 4.180V - 0.025J^- \quad (19)$$

We then compare the predicted  $\log P$  values with the fitted values and give statistics of the comparison in Table 8, where AE is the average error, AAE is the absolute average error, RMSE is the root-mean-square error, and SD is the standard deviation between the two sets of values. We also give the average  $\log P$  value for the various solvents. The results seem quite reasonable; eqs 8–13 yield descriptors that in turn predict  $\log P$  values that agree with those fitted with the descriptors in Table 6 with an average SD of 0.55 log units. Of course, this includes errors in the fitting procedure as well as errors in the predictions.

Now that it seems as though eqs 8–13 yield reasonable values of  $\log P$  for the carboxylate ions, we can use data for 13 3,4-disubstituted benzoic acids from Pytela et al.<sup>50</sup> as a test set. It is then possible to compare predictions through eqs 8–13 with experimental results that have not been used in any of the calculations. We use all the solvents studied by Pytela et al.,<sup>50</sup> except pyridine, for which we have no equation for neutral species. The experimental data are given in the Supporting Information (Table S2). The descriptors for the neutral carboxylic acids that we use as input are in Table 9, and the observed and predicted values of  $\log P$  for carboxylate anions are in Table 10. There are systematic offsets as shown by values of AE as large as 0.44 log units, but the SD values are quite good and range from 0.38 to 0.69 log units, depending on the solvent. For all 78  $\log P$

TABLE 9. Descriptors for the Benzoic Acids Used as a Test Set

acid	no.	<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>	<i>V</i>
3-methyl-4-chlorobenzoic acid	1	0.840	1.03	0.62	0.25	1.1950
3-chloro-4-nitrobenzoic acid	2	1.250	1.45	0.72	0.46	1.2283
3-methyl-4-nitrobenzoic acid	3	0.990	1.18	0.66	0.55	1.2468
3,4-dimethoxybenzoic acid	4	0.890	1.57	0.58	0.76	1.3309
3-nitro-4-methylbenzoic acid	5	0.990	1.18	0.66	0.53	1.2468
3,4-dinitrobenzoic acid	6	1.250	1.90	0.75	0.62	1.2801
3-methyl-4-methoxybenzoic acid	7	0.900	1.12	0.59	0.51	1.2722
3-methoxy-4-methylbenzoic acid	8	0.850	1.00	0.61	0.54	1.2722
3-methoxy-4-chlorobenzoic acid	9	0.940	1.19	0.64	0.47	1.2537
3-methoxy-4-nitrobenzoic acid	10	1.080	1.31	0.70	0.70	1.3055
3-bromo-4-methylbenzoic acid	11	1.000	1.10	0.70	0.27	1.2476
3-bromo-4-methoxybenzoic acid	12	1.120	1.34	0.62	0.47	1.3063
3-nitro-4-methoxybenzoic acid	13	1.120	1.49	0.66	0.65	1.3055

TABLE 10. Observed and Predicted Values of log *P* for Benzoate Anions from Water to Solvents

no. <sup>a</sup>	methanol		DMSO		DMF		propanone		acetonitrile		12DCE	
	obsd	pred	obsd	pred	obsd	pred	obsd	pred	obsd	pred	obsd	pred
1	0.06	0.44	-4.84	-4.80	-5.78	-5.74	-7.36	-7.46	-5.59	-5.80	-6.53	-6.49
2	-0.55	-0.31	-4.07	-4.77	-5.88	-6.22	-7.41	-8.15	-5.40	-6.23	-7.00	-7.14
3	-0.55	-0.61	-5.15	-5.93	-6.70	-7.11	-7.96	-8.92	-6.09	-7.08	-7.12	-7.80
4	-1.49	-1.44	-6.87	-6.45	-8.07	-7.66	-9.91	-9.59	-7.44	-7.54	-8.33	-8.25
5	-0.56	-0.53	-5.04	-5.81	-6.52	-6.99	-8.08	-8.80	-6.16	-6.97	-7.32	-7.69
6	-0.80	-1.02	-3.75	-4.84	-5.50	-6.46	-7.57	-8.56	-5.29	-6.44	-7.24	-7.45
7	-0.90	-0.38	-6.63	-5.86	-7.55	-6.90	-9.41	-8.68	-7.13	-6.88	-7.83	-7.53
8	-0.82	-0.46	-6.42	-6.25	-7.49	-7.24	-9.10	-8.98	-7.02	-7.22	-7.81	-7.82
9	-0.50	-0.29	-5.49	-5.51	-6.60	-6.62	-8.17	-8.44	-6.27	-6.63	-7.10	-7.34
10	-1.96	-1.04	-5.50	-6.30	-6.75	-7.61	-8.45	-9.52	-6.46	-7.58	-7.66	-8.34
11	-0.42	0.58	-5.01	-4.47	-6.17	-5.53	-8.04	-7.37	-6.13	-5.70	-7.10	-6.45
12	-0.43	-0.11	-5.14	-4.98	-6.37	-6.21	-8.12	-8.13	-6.13	-6.26	-6.89	-7.02
13	-1.07	-0.92	-5.41	-5.74	-6.91	-7.10	-8.62	-9.05	-6.45	-7.06	-7.65	-7.86
AE	-0.30		0.18		0.09		0.26		0.44		0.12	
AAE	0.34		0.51		0.41		0.55		0.55		0.29	
RMSE	0.45		0.60		0.50		0.65		0.67		0.37	
SD	0.47		0.63		0.52		0.67		0.69		0.38	

<sup>a</sup>Numbering as in Table 9.

values, AE = 0.13, AAE = 0.44, RMSE = 0.54, and SD = 0.55 log units, and we can take these as estimates of the predictive ability of our method. The predictions require as input only the descriptors of the corresponding neutral carboxylic acids, and can be made for any solvent for which  $J^-$  is known, see Table 3. Whether or not predictions could (or should) be amended to take into account the offset AE values for particular solvents is a moot point.

The carboxylic anions that we use in the training set and in the test set contain a variety of functional groups, so that the equations we set up for predictions seem to be very general. The only functional group that we have deliberately omitted is the carboxylate group itself, as in dicarboxylic acids. We deal with these compounds separately.

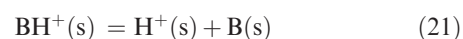
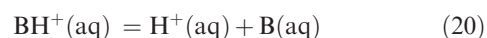
There are a number of solvents that we have not considered. For some of these, such as pyridine and sulfolane, we lack the appropriate equation for the neutral compounds. For others, such as wet octanol, nitrobenzene, and tetrahydrofuran, we need a more detailed analysis that we hope to provide in the future.

**Calculations for Carboxylate Dianions.** We were interested to see what the descriptor values were for dianions of carboxylic acids, but as a preliminary step we first studied the monoanions derived from dicarboxylate acids. Descriptors for the monoanions and the dianions from dicarboxylic acids are shown in Table 11. The latter descriptors can only be approximate because they require both the first and the second acid

dissociation constants in water and the nonaqueous solvents. Details of the  $pK_a$  values, with references, are shown in Table S2 (Supporting Information).

The descriptors for the monoanions are of some interest in that the  $A_i$  values are usually far less than expected for solutes that carry the carboxylic acid group. The  $-\text{CO}_2^-$  group must be a very powerful electron-donating group that considerably reduces the hydrogen bond acidity of the remaining  $-\text{CO}_2\text{H}$  group. The aliphatic monoanions have nearly zero  $A_i$  values, and the only real exception is the monoanion of phthalic acid. Interestingly, the  $S_i$  value for the phthalate and the maleate monoions are much smaller than those for the isophthalate and fumarate monoions. This may be due to intramolecular hydrogen bonding between the  $\text{C}-\text{O}-\text{H}$  and the  $\text{C}-\text{O}^-$  groups held in the cis-position, even though a seven-membered ring has to be formed. All the dianions are characterized by very large  $S_i$  and  $A_i$  values and so are among the most dipolar and hydrogen bond basic solutes we have studied.

**Calculations for Protonated Amines.** The equations needed to obtain log  $P$  values from  $pK_a(\text{aq})$ ,  $pK_a(\text{s})$ , and log  $P(\text{H}^+)$  are somewhat different from those for the carboxylate anions because the protonated amines are now the reactant in eq 20 rather than one of the products.



**TABLE 11. Descriptors for Monoanions and Dianions of Dicarboxylic Acids Obtained Using the  $pK_a$  Method**

anion	$E_i$	$S_i$	$A_i$	$B_i$	$V_i$	$J^-$	$N$	SD
phthalic acid, monoanion	1.25	1.58	0.63	2.03	1.1255	0.672	8	0.363
isophthalic acid, monoanion	1.25	4.03	0.26	2.95	1.1255	2.285	7	0.249
terephthalic acid, monoanion	1.25	3.90	0.41	2.92	1.1255	2.323	7	0.138
malonate, monoanion	0.53	1.15	0.11	1.93	0.6586	0.573	6	0.173
succinic acid, monoanion	0.52	1.03	0.00	2.35	0.7995	0.783	7	0.086
maleic acid, monoanion	0.86	0.97	0.00	1.91	0.7565	0.625	7	0.417
fumaric acid, monoanion	0.86	3.45	0.08	2.85	0.7565	2.087	5	0.094
phthalic acid, dianion	1.40	3.88	0.00	5.25	1.1040	3.124	4	0.465
isophthalic acid, dianion	1.40	3.36	0.00	5.22	1.1040	3.216	4	0.421
terephthalic acid, dianion	1.40	3.55	0.00	5.08	1.1040	3.356	3	n/a
malonic acid, dianion	0.68	4.88	0.00	6.22	0.6371	4.523	4	0.146
succinic acid, dianion	0.67	4.21	0.00	6.24	0.7780	4.437	4	0.238
maleic acid, dianion	1.01	3.82	0.00	5.88	0.7350	4.000	4	0.560
fumaric acid, dianion	1.01	5.29	0.00	5.46	0.7350	4.006	4	0.047

For transfer from water to the other phase,  $\log P(\text{BH}^+)$  is given by eq 22

$$\log P(\text{BH}^+) = \log P(\text{B}) + \log P(\text{H}^+) - pK_a(\text{aq}) + pK_a(\text{s}) \quad (22)$$

There is much less data on  $pK_a(\text{s})$  values for protonated amines than for the carboxylic acids. The comprehensive review on cations by Kalidas et al.<sup>29</sup> lists only the transfer of  $\text{NH}_4^+$  from water to methanol. The review of Izutsu<sup>49</sup> gives  $pK_a(\text{s})$  for a variety of protonated amines in aprotic solvents, and values in methanol are given by Rived et al.<sup>37,51</sup> Rööm et al.<sup>52</sup> deal mainly with diamines but give values of  $pK_a(\text{s})$  for a few amines in acetonitrile and tetrahydrofuran. Zielinska et al.<sup>53</sup> have studied the butylammonium cation in a number of solvents, and Bell<sup>54</sup> lists a number of  $pK_a(\text{s})$  values in methanol and ethanol. In addition to the lack of data on  $pK_a(\text{s})$  values, we were also restricted by the number of solvents we could study due to lack of  $j^+$  coefficients; see Table 3. For the protonated amines, we adopted a converse procedure to that for the carboxylic acids by subtracting 0.15 from the  $E$  value of the neutral amine. The McGowan volume was calculated from the molecular formula of the ionized form, equivalent to addition of 0.0215 to  $V$  for the neutral amine. The descriptors for the 26 neutral amines that we used as a training set are in Table 12, and the obtained descriptors for the 26 protonated amines are in Table 13. The data that we used, together with references, are included in Table S3 (Supporting Information).

There are just enough protonated amines in the training set to obtain equations that can be used for the prediction of descriptors. As before, we used descriptors for the neutral amines as the independent variables, together with the independent variable  $\text{NA}_i$ , the number of hydrogen atoms attached to the charged nitrogen atom. Thus, for methylamine and 1,6-hexanediamine,  $\text{NA}_i = 3$  and for pyridine  $\text{NA}_i = 1$ . The full set of equations is as follows; note that  $B_i$  is

**TABLE 12. Descriptors for the Amines Used as a Training Set**

base	$E$	$S$	$A$	$B$	$V$
aniline	0.955	0.96	0.26	0.41	0.8162
3-methylaniline	0.946	0.95	0.23	0.45	0.9571
4-methylaniline	0.923	0.95	0.23	0.45	0.9571
2-chloroaniline	1.033	0.92	0.25	0.31	0.9390
4-chloroaniline	1.060	1.13	0.30	0.31	0.9390
2-nitroaniline	1.180	1.37	0.30	0.36	0.9904
3-nitroaniline	1.200	1.71	0.40	0.35	0.9904
<i>N</i> -methylaniline	0.948	0.90	0.17	0.43	0.9571
ammonia	0.139	0.39	0.16	0.56	0.2084
ethylamine	0.236	0.35	0.16	0.61	0.4902
propylamine	0.225	0.35	0.16	0.61	0.6311
butylamine	0.224	0.35	0.16	0.61	0.7720
dimethylamine	0.189	0.30	0.08	0.66	0.4902
diethylamine	0.154	0.30	0.08	0.69	0.7720
trimethylamine	0.140	0.20	0.00	0.67	0.6311
triethylamine	0.101	0.15	0.00	0.79	1.0538
pyridine	0.631	0.84	0.00	0.52	0.6753
2-methylpyridine	0.598	0.75	0.00	0.58	0.8162
4-methylpyridine	0.630	0.82	0.00	0.54	0.8162
4-chloropyridine	0.740	0.85	0.00	0.40	0.7977
4-bromopyridine	0.900	0.93	0.00	0.38	0.8503
4-pyridinecarboxaldehyde	0.796	1.12	0.00	0.74	0.8319
ethyl isonicotinate	0.756	1.05	0.00	0.71	1.1724
piperidine	0.422	0.40	0.06	0.77	0.8043
benzimidazole	1.270	1.40	0.38	0.76	0.9053
imidazole	0.710	0.85	0.42	0.78	0.5363

zero for all the protonated amines studied.

$$E_i = -0.15 + 1.00E \quad (23)$$

$$S_i = 0.463 + 0.473S + 2.419B \quad (24)$$

$$A_i = -0.052 - 0.350E + 1.480S + 0.327\text{NA}_i \quad (25)$$

$$V_i = 0.0215 + 1.00V \quad (26)$$

$$J^+ = 0.628 + 1.002E - 0.794S + 1.128B - 0.191\text{NA}_i \quad (27)$$

We can now use eqs 23–27 to predict descriptors for protonated amines as a test set, that is, compounds that have not been used in any way to obtain eqs 23–27. Since it is not possible to compare the predicted descriptors with “observed” descriptors, we use the predicted descriptors to predict  $\log P$  values for transfer of 18 protonated amines from water to various other solvents. Details are shown in Table 14.

(51) Rived, F.; Canals, I.; Bosch, E.; Roses, M. *Anal. Chim. Acta* **2001**, *439*, 315–333.

(52) Rööm, E.-I.; Kutt, A.; Kaljurand, I.; Koppel, I.; Leito, I.; Koppel, I. A.; Mishima, M.; Goto, K.; Miyahara, Y. *Chem.—Eur. J.* **2007**, *13*, 7631–7643.

(53) Zielinska, J.; Makowski, M.; Maj, K.; Liwo, A.; Chmurzynski, L. *Anal. Chim. Acta* **1999**, *401*, 317–321.

(54) Bell, R. P. *The Proton in Chemistry*; Methuen & Co. Ltd.: London, 1959.



TABLE 13. Descriptors for Protonated Amines in the Training Set Obtained Using the  $pK_a$  Method<sup>a</sup>

base	$E_i$	$S_i$	$A_i$	$V_i$	J+	NA <sub>i</sub>	$N^b$	SD <sup>b</sup>
aniline	0.805	1.62	1.93	0.8377	0.620	3	6	0.425
3-methylaniline	0.796	1.92	2.05	0.9786	0.750	3	3	0.322
4-methylaniline	0.773	1.99	1.98	0.9786	0.668	3	5	0.311
2-chloroaniline	0.883	1.70	1.93	0.9605	0.730	3	1	n/a
4-chloroaniline	0.910	1.90	2.40	0.9605	0.633	3	3	0.192
2-nitroaniline	1.030	2.40	2.55	1.0119	0.690	3	2	0.098
3-nitroaniline	1.050	1.94	3.06	1.0119	0.076	3	5	0.058
<i>N</i> -methylaniline	0.798	1.87	1.53	0.9786	0.892	2	3	0.003
ammonia	-0.011	1.77	1.80	0.2299	0.370	4	6	0.162
methylamine	0.100	2.90	1.35	0.3708	0.722	3	3	0.006
ethylamine	0.086	2.50	1.31	0.5117	0.768	3	4	0.281
propylamine	0.075	2.18	1.40	0.6526	0.675	3	3	0.029
butylamine	0.074	1.93	1.30	0.7935	0.817	3	4	0.213
dimethylamine	0.039	2.41	1.00	0.5117	0.877	2	3	0.002
diethylamine	0.004	2.30	0.98	0.7935	0.930	2	2	0.247
trimethylamine	-0.010	2.19	0.55	0.6526	1.064	1	3	0.237
triethylamine	-0.049	2.20	0.60	1.0753	1.250	1	3	0.216
pyridine	0.481	2.46	1.15	0.6968	1.050	1	4	0.444
2-methylpyridine	0.448	2.22	1.19	0.8377	1.120	1	2	0.199
4-methylpyridine	0.480	2.00	1.28	0.8377	1.190	1	1	n/a
4-chloropyridine	0.590	1.65	1.29	0.8192	0.980	1	1	n/a
4-bromopyridine	0.750	1.70	1.36	0.8718	0.980	1	1	n/a
4-pyridinecarboxaldehyde	0.646	2.97	1.60	0.8534	1.180	1	1	n/a
ethyl isonicotinate	0.606	2.55	1.60	1.1939	1.220	1	1	n/a
piperidine	0.272	2.73	0.98	0.8258	1.178	2	4	0.260
benzimidazole	1.120	3.00	2.23	0.9268	1.370	2	1	n/a
imidazole	0.560	2.50	1.61	0.5578	1.050	2	1	n/a

<sup>a</sup>NA<sub>i</sub> is the number of hydrogen atoms attached to the charged nitrogen atom. <sup>b</sup> $N$  is the number of solvents (log  $P$  values) used to obtain the descriptors, and SD is the standard deviation in calculated and observed log  $P$  values.

TABLE 14. Predicted Descriptors and Observed and Predicted Values of log  $P$  for Transfer of Protonated Bases from Water to Solvents

base	$E_i$	$S_i$	$A_i$	$V_i$	J+	NA <sub>i</sub>	solvent	log $P$	
								obsd	calcd
3-chloroaniline	0.903	1.71	2.19	0.9605	0.575	3	methanol	1.44	1.80
3-bromoaniline	0.978	1.75	2.30	1.0127	0.579	3	methanol	1.44	2.00
4-nitroaniline	1.070	2.22	3.34	1.0119	0.148	3	propanone	2.38	2.00
4-hydroxyaniline	1.000	2.92	2.70	0.8964	0.864	3	methanol	0.39	0.01
3-hydroxyaniline	0.980	2.62	2.96	0.8964	0.528	3	methanol	0.80	1.16
isobutylamine	0.048	2.14	1.33	0.7935	0.710	3	MeCN	-0.40	0.04
<i>tert</i> -butylamine	-0.029	2.32	1.32	0.7935	0.747	3	MeCN	-0.64	0.02
hexylamine	0.047	2.10	1.38	1.0753	0.663	3	methanol	1.03	1.21
dipropylamine	-0.026	2.27	1.00	1.0753	0.910	2	DMSO	3.11	3.48
pyrrolidine	0.256	1.93	1.87	0.6849	0.372	2	DMSO	4.54	4.85
2-aminoethanol	0.308	2.96	1.76	0.5704	0.997	3	methanol	-1.40	-1.88
2-aminoethanol	0.308	2.96	1.76	0.5704	0.997	3	MeCN	-2.07	-1.77
1,6-hexanediamine	0.266	3.55	1.73	1.1751	1.261	3	methanol	-0.50	-0.87
1,8-octanediamine	0.262	3.57	1.79	1.4569	1.225	3	methanol	0.44	0.22
1,10-decanediamine	0.260	3.58	1.82	1.7387	1.207	3	methanol	1.46	1.27
1,11-undecanediamine	0.259	3.58	1.82	1.8796	1.206	3	methanol	1.95	1.77
1,12-dodecanediamine	0.258	3.58	1.82	2.0205	1.205	3	methanol	2.46	2.27
urea	0.351	3.20	2.96	0.4861	0.321	3	MeCN	-2.74	-2.27
morphine	1.970	5.95	1.25	2.0863	3.941	1	methanol	0.51	-5.88
morphine	1.970	5.95	1.25	2.0863	3.941	1	propanone	-2.03	-1.36
diphenylamine	1.320	1.76	2.04	1.4455	0.987	2	propanone	5.17	2.71
codeine	1.810	5.72	2.47	2.2272	2.872	1	propanone	-0.41	0.89

For 18 protonated amines in Table 14, the predicted values of the ionic descriptors yield predicted values of log  $P$  from water to various solvents that can be compared to the observed values. For all 18 log  $P$  values, AE = -0.090, AAE = 0.355, RMSE = 0.379, and SD = 0.390 log unit. Although these statistics provide an estimate of the ability of equations to predict log  $P$  values, it must be noted that both the training set and the test set are restricted in terms of the type of amine base. We were unable to predict satisfactorily log  $P$  values for protonated morphine, codeine, or diphenyl-

amine, possibly because the bases were outside the chemical space of the training set. In addition, there can be serious errors in the observed  $pK_a$ (s) values. There is no other way to interpret the reasonable fit for protonated morphine for transfer to propanone and the extraordinarily poor fit for transfer to methanol. It is, unfortunately, not uncommon for there to be serious discrepancies in recorded data. For example  $pK_a$ (s) for *N,N*-dimethylaniline in propanone is given as 4.91<sup>49</sup> and 7.24<sup>36</sup> and for *N,N*-diethylaniline in propanone values are 6.26<sup>49</sup> and 9.17,<sup>36</sup> so that we were

**TABLE 15.** Comparison of Descriptors for Protonated Amines, with Those for Other Compounds

compd	<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>
propylamineH <sup>+</sup>	0.075	2.18	1.40	0.00
dimethylamineH <sup>+</sup>	0.039	2.41	1.00	0.00
triethylamineH <sup>+</sup>	-0.049	2.20	0.60	0.00
4-methylanilineH <sup>+</sup>	0.773	1.99	1.98	0.00
3-nitroanilineH <sup>+</sup>	1.050	1.94	3.06	0.00
pyridineH <sup>+</sup>	0.481	2.46	1.15	0.00
acetic acid	0.265	0.64	0.62	0.44
trichloroacetic acid	0.524	1.21	1.01	0.26
ethanol	0.246	0.42	0.37	0.48
benzoic acid	0.730	0.90	0.59	0.40
2,4-dichlorobenzoic acid	0.950	1.00	0.75	0.37
phenol	0.805	0.89	0.60	0.30
3,5-dinitrophenol	1.320	2.18	1.05	0.16

unable to use either of these amines. There are many other such instances, which is one of the reasons why we have comparatively little data on the amine bases in Tables 13 and 14.

There is, however, enough data to compare the ionic descriptors for protonated amines with those for neutral species. Values of *E*, *S*, *A*, and *B* are in Table 15 for typical protonated amines and for a selection of neutral solutes. Just as the carboxylate anions are the strongest monofunctional hydrogen bond bases, so the protonated amines, especially those derived from anilines, are the strongest monofunctional hydrogen bond acids.

We had hoped to obtain descriptors for diprotonated amines, but were prevented due to lack of the required  $pK_a(s)$  values. Judging from our work on the dianions of dicarboxylic acids, there is no reason in principle why descriptors for dications of diamines could not be obtained.

As regards application of the ionic descriptors to any particular process, a specific requirement is a knowledge of

the coefficients  $j^+$  and  $j^-$  in eq 2. Unless  $j^+$  and  $j^-$  can be predicted in some way, values of the dependent variable for a few ions or ionic species are needed in order to determine  $j^+$  and  $j^-$ . Once these coefficients are known, then the dependent variable can be predicted for all the ionic species for which we have descriptors.

## Conclusions

We have used a new method for the determination of partition coefficients for ionic species based on  $pK_a$  values in water and nonaqueous solvents. From the obtained partition coefficients, descriptors have been determined for 71 carboxylate anions and equations set up for the prediction of descriptors for further carboxylate anions. These anions are the most powerful hydrogen bond monofunctional bases we have studied, being much more powerful than neutral solutes such as tributylphosphine oxide or triethylamine. There is less data for protonated amines, but we have assembled descriptors for 26 protonated amines and constructed equations that we have used to predict descriptors for 18 protonated amines in a test set. The protonated amines, especially protonated anilines, are very powerful monofunctional hydrogen bond acids: protonated anilines are much stronger hydrogen bond acids than trichloroacetic acid.

**Acknowledgment.** We are grateful to Martí Rosés, Clara Ràfols, Elisabeth Bosch, and Joaquim Sales for their help.

**Supporting Information Available:** Table S1 containing partition coefficients for ions and Tables S2 and S3 containing details of data used for calculations of carboxylate anions and protonated amines. This material is available free of charge via the Internet at <http://pubs.acs.org>.